DOCUMENT RESUME

ED 351 374 TM 019 198

TITLE	Tests, Examinations and Assessment in Schools: The Role of External Tests and Examinations in the Assessment of Student Achievement at the Secondary Level--A Review of the Current Debate.
INSTITUTION	Northern Territory Dept. of Education, Darwin (Australia).
REPORT NO	CAB-R-2/1988; ISBN-0-7245-1991-2
PUB DATE	89
NOTE	44p.
PUB TYPE	Information Analyses (070) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; *Achievement Tests; Comparative Analysis; *Educational Assessment; Educational Change; Educational Improvement; Educational Practices; Educational Trends; Foreign Countries; Secondary Education; *Secondary School Students; *Standardized Tests; *Student Evaluation; *Test Use
IDENTIFIERS	*Australia (Northern Territory); External Examination Program

ABSTRACT

Issues surrounding educational assessment in Australia's Northern Territory are discussed. It is possible that the real issue is not an actual decline in educational standards, but instead, a decrease in public confidence in educational practice. Global patterns in educational testing and the assessment systems of the United States, the British Isles, and parts of Europe and Asia are compared with those of the individual Australian states and territories. Arguments against system-wide external examinations and tests for the Northern Territory can be summarized as: (1) in the rest of Australia, the trend is away from such examinations; (2) these tests may encourage classroom competition rather than cooperation; (3) failure may become institutionalized; (4) reform may be frustrated by external examinations; (5) inequalities may be supported by test results; (6) tests do not necessarily work to raise standards; (7) tests deny the realities of individual differences; (8) the validity of tests is in question; and (9) improper test use can give inaccurate portrayals of achievement. Proponents of standardized assessment argue that objective system-wide tests negate the potential prejudices of subjective assessment and facilitate advancement of minority groups. Possible approaches that the Territory may choose are outlined. A 68-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED351374

CURRICULUM AND ASSESSMENT BRANCH

RESEARCH AND EVALUATION REPORTS

TM019198

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Cliff Fowler

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Curriculum and Assessment Branch

Research and Evaluation Report Number 2/1988

TESTS, EXAMINATIONS AND ASSESSMENT IN SCHOOLS:

The Role of External Tests and Examinations in the Assessment of Student Achievement at the Secondary Level - a Review of the Current Debate

Northern Territory Department of Education
Darwin 1989

© Northern Territory Department of Education 1989

ISBN 0 7245 1991 2

For further information please contact:
PEO Special Projects
Curriculum & Assessment Branch
GPO Box 4821 Darwin NT 0801
Telephone: (089) 89 5611

Printed by the Distance Education Branch of
the Northern Territory Department of Education
P&PMarch89/360—450

Stores Stock Number 810—0018

Acknowledgments

The author wishes to acknowledge the assistance of Mr C F Fowler Superintendent, Research and Evaluation Section, Curriculum and Assessment Branch; and Dr C H Payne Director, Curriculum and Assessment Branch, Northern Territory Department of Education in the preparation and publication of this report.

The assistance received from the officers of other State and Territory Departments of Education, statutory boards and authorities in the preparation of this report is also acknowledged.

Dr Richard Watkins, Research and Evaluation Section, Curriculum and Assessment Branch, Northern Territory Department of Education, Darwin, December 1988.

Watkins, Richard.
Tests, examination and assessment in schools : the role of external tests and examinations in the assessment of student achievement at the secondary level : a review of the current debate / (Richard Watkins)
Darwin : Northern Territory Dept. of Education, 1989.
iii, 39 p. : 30 cm. (Research and evaluation report : no. 2/1988.)
Bibliography: p. 36-39.
ISBN 0 7245 1991 2
1. Educational tests and measurements. I. Northern Territory. Dept. of Education. II. Northern Territory. Curriculum and Assessment Branch.
III. Title. IV. Series: Research and evaluation report (Northern Territory. Curriculum and Assessment Branch) : no. 2/1988.
371.26 19

Table of Contents

Chapter 1 - Introduction and background	page	1
Evidence of concern about education and student attainment	page	1
Chapter 2 - Literacy and Numeracy	page	5
Consideration of terms and meanings	page	5
Chapter 3 - The Use of External Examinations Tests	page	8
Global patterns	page	8
Australian States and Territories	page	12
Chapter 4 - Arguments about External Examinations and Tests	page	16
Precursors to the debate	page	16
System-wide Examinations and Tests: a Case Against	page	17
System-wide Examinations and Tests: a Case For	page	23
Chapter 5 - Possible ways forward	page	33
Responses	page	33
Bibliography and References	page	36

Chapter 1

Evidence of concern about education and student attainment

In Towards the 90s: Volumes 1 and 2, it is stated that: "The Department of Education has a responsibility not only to continually strive to raise the standards of literacy and numeracy of the Northern Territory school population, but also to increase public confidence in the results and qualifications achieved by students in schools' programmes." (Volume 1, page 9)

The above statement reflects a number of the wider and more generalised concerns present in Australia and elsewhere, about the maintenance and improvement of standards in schools, particularly at the conclusion of compulsory education which generally ends with the completion of Year 10 in the Northern Territory. Departmental, public and political concern about enhancing and improving education is by no means new, nor are widespread concerns about the impact of changes in the means by which student performance and attainment are monitored. Moreover, current education policy both within and outside the Northern Territory, is seeking to address critical aspects of student performance and attainment, with a view to optimising levels of achievement and avoiding what has been termed the 'self-fulfilling prophecy' which holds that students will tend to do as well as expected, rather than as well as they are capable.

It should be noted that statements made by the Minister of Education, and by others, do not suggest that standards in literacy and numeracy have declined, or that there is a causal relationship between standards and tests or examinations. What is being suggested is a very clear concern both locally and nationally to bring about improvements in present standards, so that the Australian community is better able to adapt to change, more highly skilled and competitive on a global basis.

These concerns were touched upon by the Minister for Employment, Education and Training in the recent paper Strengthening Australia's Schools (Canberra, 1988), in which it states:

"Schools play a critical and central role in the nature of our society and economy, and much is expected and demanded of them. We need to ensure that all Australian schools are of the highest possible standard."; and "We must continually look for ways to improve the quality, relevance and effectiveness of schools throughout Australia." (foreword)

Strengthening Australia's Schools continues:

"We need to ensure that every young Australian gets a general education of quality which provides both personal and intellectual development as well as broadly based and adaptable skills.". (page 2)

Further: "The issue here is not the level of our investment as a nation in our schools but rather the quality and appropriateness of their achievements to those young Australians who pass through them.". (page 2)

In an earlier paper titled Skills for Australia, (Canberra, 1987) it is stated:

"First there is a need to boost the general level of education and training within the Australian community, in quality and quantity.". (page 6) and; "A high quality basic education is an essential pre-requisite for a vocationally skilled and adaptable labour force. More needs to be known about the levels of competence achieved by our students at school, especially in the core disciplines of language, mathematics and science.". (pages 8 and 9)

In the Beazley Report (1984), such concerns were expressed in the following way:

"It was accepted that the maintenance of educational standards is a top educational priority ...." (page 133).

Beazley also reports: "The Committee received many written submissions referring to standards in literacy and numeracy, and received much oral evidence on the subject. Rightly or wrongly, there is a growing belief that standards are slipping below acceptable levels and that many students are poorly educated and even unemployable." and;"In responding to this concern, the Committee did not start from the position that standards of literacy and numeracy and written and spoken expression are declining, for it has seen no clear research evidence supporting that contention. The Committee started from the position that community and business life requires higher standards than those of the past and if students cannot read properly, cannot spell accurately, cannot write grammatically correct sentences, cannot calculate or estimate, they are severely disadvantaged.". (page 29)

More recently, evidence has emerged from both Britain and the United States about considerable public disquiet over the 'state' of current education, with nation-wide moves to implement or re-introduce national or state based objective tests and examinations, and to focus much more strongly on clearly defined curriculum and educational goals.

In the United States it is reported that: "The move towards state-mandated testing has been strongest in the southern states where educational spending and performance have traditionally been the lowest - but which are now starting to rise in the schooling league.". (The Economist 31 October 1987)

General Australian concerns about education, as expressed in Beazley (ibid), continue, for example with a contemporary editorial in The Australian, entitled "The Way Ahead IV", suggesting that:

"There is literally nothing on the national agenda that cries out more urgently for attention than the need for major educational reform. ...."

"Education in Australia is nothing short of a disaster and the most radical and urgent action is needed."

".... In education our aims have become confused and diluted, our performance has slipped as a consequence, and we do not even have the means to monitor our performance. ...."

"It is futile cowardly and ultimately self-defeating for schools to attempt to abolish competition. ...."

Further: "It is against common sense and it is against international experience to suggest that high standards can be maintained without objective testing of those standards and without a competitive approach to achievement.". (The Australian, 6 January 1988)

Numerous subsequent letters to this newspaper reflected widespread agreement with the sentiments expressed, with many calls for the re-introduction of examinations, clearly defined curriculum objectives and a 'back to basics' approach in respect of literacy, numeracy and the teaching of essential subjects such as English, Mathematics, Science and Social Studies.

In a later editorial entitled "Language the bricks, mortar of knowledge", The Australian wrote:

"One of the most important aspects of education is the teaching of the English language. The anecdotal evidence from employers of a level of functional illiteracy among job applicants continues to grow .... " (The Australian, 14 January 1988)

From the above article, the editorial developed arguments in favour of a return to the teaching of literature and formal grammars which would enable young people to deal effectively with the complex society which surrounds them.

The debate about the standards of literacy and numeracy in Australia was renewed in late 1988 following a statement by Leonie Kramer in the 'Hotline' segment of The Australian newspaper (2 December 1988). During the week in which 'Hotline' ran, The Australian received a total of 1,045 telephone calls about the literacy topic. Of these calls, 89 per cent were reported as 'for' the main tenor of Kramer's argument that literacy should be more heavily emphasised in schools and the wider community, 4 per cent were reported as 'against', and 7 per cent reported as 'others'.

In the local press, The NT News editorial of 28 January 1988 under the title "Malaise in Education", which is symptomatic of many recently occurring head-lines and newspaper articles, wrote:

"The Northern Territory shares with the States a profound malaise in education. Too many of our young leave school functionally illiterate and innumerate, unable to go into apprenticeships or to enter the work-force."

"It is fair for the average citizen to ask what is being achieved for that enormous sum of money ($213 million) and what precisely are the education goals in the Northern Territory."

"It is no coincidence that nations with educational systems that place heavy emphasis on literacy and numeracy skills in the early years also have far higher proportions of their populations in advanced education."(NT News, 28 January 1988)

It is clear that the quotations and statements above frequently do no more than present quite broad assertions about the 'condition' of education. This may not be unreasonable given that they are essentially editorials rather than attempting to be definitive analyses of national or state positions vis a vis education.

What is present in most of the many articles, and in many other criticisms of education systems, is an often vividly expressed concern that young Australians (and young Britons and Americans) are finishing their compulsory education with what are described as inadequate standards of literacy and numeracy. While, literacy and numeracy are left generally undefined, the tenor of the arguments appears to revolve around the apparent inability of young school leavers to function effectively in contemporary society.

The examples most often employed to illustrate this phenomena are: very poor levels of achievement shown in job applications; high failure rates among apprenticeship and other technical trainees; poor reading and comprehension skills; inabilities in handling money and basic arithmetic tasks; and weak verbal communication skills. Perhaps too simplistically, perceptions of the underlying causes for these 'problems' are held to be in English and Mathematics teaching, the syllabuses and curriculum for these subjects and other general inadequacies in contemporary education systems.

It may be argued that these perceptions lack the sophistication necessary to address complex questions of social and educational behaviour, however they do represent apparently quite widespread public concerns and therefore should not fail to be addressed. Further, if these expressed public concerns are directed inappropriately at English and Mathematics per se, it becomes incumbent upon education authorities to, on the one hand, provide countervailing evidence that teaching and learning within these disciplines is appropriate and at the very least, adequate; and secondly to provide comprehensive information about the extent to which young people have mastered aspects of these subjects which are considered 'essential', *eg* a variety of writing, speaking and reading styles, and arithmetical skills.

It may be that the provision of system-wide information in a publicly comprehendable form which appears to be currently absent from educational reporting, is partially a root cause of current concerns. It may also be that system-wide objective analysis of 'essential' elements of literacy and numeracy, which are incorporated in secondary English and Mathematics syllabuses, will satisfy public demands for educational accountability.

Finally, it is important to note that the media rhetoric of claim and counter-claim regarding a perceived decline in educational standards deals largely in anecdote and personal opinion. As the Beazley Report (ibid) recognised, there is "..... no clear research evidence supporting that contention". Indeed, the periodic surveys of literacy and numeracy levels of primary students conducted by the Queensland Education Department over two decades show no decline in standards. Similarly, the recently published 1987 Survey of Basic Reading Skills of 14-Year-Old Tasmanian Students (Education Department, Tasmania, November, 1988) shows a steady improvement from 1975 to 1983, and no statistically significant change from 1983 to 1987. At the tertiary level, the research reported by Gordon Taylor (1988) showed no significant difference in the ability to handle the English language of first year students at Monash University between 1974 and 1984.

It appears therefore that the real issue may well be related, not to any actual decline in educational standards, but rather to a loss in public confidence in current educational practice as it is perceived by the wider community.

Chapter 2

Literacy and Numeracy

Consideration of terms and meanings

It appears useful that questions pertaining to literacy and numeracy are addressed, and that adequate 'working' definitions are presented. Further, literacy and numeracy should be placed in the context of English and Mathematics as curriculum areas and subjects, and the meanings of the former should be clarified to avoid possible confusion and inter-change in the terms.

In recent editorials and newspaper articles, and in other sources, consistent reference has been made to literacy and numeracy and to English and Mathematics as subject or curriculum areas. Often these terms have been used inter-changeably and as synonyms for one another, with no attempt to clarify what the authors intend. Indeed, a valid criticism of Towards the 90s Volume 1, is that it too employs the terms literacy and numeracy, in the context of English and Mathematics as curriculum areas, but provides only inferential references in respect of what is intended.

Attempts to provide a precise definition of literacy and numeracy in relation to education, will almost certainly founder because of the inherent complexities which attend polysemic expressions of this type. It is for this reason that a 'working' definition appears to offer one way forward, and a method by which wider audiences may grasp the intent of the terms, and place them in a context of English and Mathematics as subject or curriculum areas, together with other subjects and curriculum typically offered in secondary education..

As a means of illustrating the complexities of meaning which attend the terms literacy and numeracy, it may be appropriate to consider the following:

Literacy - ability to read and write. (Concise Oxford Dictionary,1976)

Numeracy - (cf numerate) - acquainted with the basic principles of mathematics and science.(Concise Oxford Dictionary,1976)

Literacy - (cf literate) - learned; able to read and write. (Concise Oxford Dictionary,1976)

Numeracy - (cf numerate) - having some understanding of mathematics and science. (Chambers 20th Century Dictionary, 1983)

Literacy (cf literate) - primed, acquainted, informed, lettered, schooled, learned, well-informed, knowledgeable, scholastic, cultured, civilised, sophisticated. (Chambers 20th Century Dictionary, 1983)

Many persons within the wider community may find such dictionary based definitions acceptable within a context of general, rather than particular circumstances, and when used conversationally to express an overall point of view. On the other hand, when placed in the context of education, and juxtaposed against measures of student performance and attainment, such definitions might appear unduly superficial and potentially misleading in that they fail to address the **specific** elements of what may constitute literacy and numeracy.

One possible solution to the dilemma of a working definition may be found in the Beazley Report (ibid), which states:

"A person is considered to be literate and numerate when he has acquired the skills and concepts which enable him to function effectively in his group and community, and when his attainments in reading writing and mathematics make it possible for him to continue to use these skills to further his own and his community's development." (page 123)

While this 'definition' does not itself provide a precise or quantitative answer, it does suggest a way forward which can itself be placed in the context of additional qualifying statements which address the question of **specific** elements noted earlier (cf Education Department of Western Australia 1977, The College Board 1983; and Mathematics Counts, 1982).

Examples of the types of specific category which might be found within literacy are:

reading - including identifying and comprehending main and subordinate ideas; recognition of the different purposes of writing; separation of personal and authorial ideas; establishing variation in reading speed and method; use of references and sources of information; and ability to define unfamiliar words and phrases by contextual reference;

writing - an ability to: conceive ideas about a topic; write standard English using correct grammatical structure; vary individual writing style to suit different readers and purposes; improve writing by the correction of stylistic and other errors; and gather information from primary and secondary sources and prepare reports, summaries and paraphrasing using this data; and

speaking and listening - an ability to: engage critically and constructively in exchanges of ideas; to ask and respond to questions and follow spoken instructions; identify and comprehend main and subordinate ideas and report accurately what others have said; and conceive and develop ideas about a topic for the purposes of public speaking.

Examples of the types of specific category which might be found within numeracy are:

mathematics - an ability to undertake: computations including addition, subtraction, multiplication and division using natural numbers, fractions, decimals and integers; measurement in metric units; effective use of - integers, fractions and decimals; ratios, proportions and percentages; roots and powers; algebra; and geometry;

number - students should possess a reliable method of carrying out calculations, and be encouraged to develop intuitive approaches to written and mental calculations;

money - recognise coins and notes and apprehend sub-divisions of currency into smaller units; and

other concepts including: time, measurement, graphs and pictorial representations, spatial reasoning, and statistical ideas.

It is possible to view literacy and numeracy as phenomena which span the whole of the contemporary curriculum, and not solely English and Mathematics.

It appears equally possible to argue that the actual 'content' of numeracy and literacy in their broadest sense, is encompassed within the typical school curriculum, if not specifically within the English and Mathematics curriculum per se, then certainly within the broad ambit of the subjects a student would normally expect to encounter during secondary education. Indeed, it is difficult to imagine any contemporary curriculum area, or typical secondary subject in which teaching and learning would not contain many of the items listed above on a regular and frequent basis, and which failed to address these as essential components of student learning.

Within the Northern Territory, which has a Board Approved Curriculum and Core Objectives, students may expect to encounter many if not all the areas which are typically encompassed by literacy and numeracy in fields such as: business education; technical studies; social and cultural education and science. Moreover, it would be remiss to suggest that teachers outside English and Mathematics ignore these areas, leaving them to subject specialists within 'relevant' or specialised disciplines.

Because English and Mathematics as subjects or disciplines contain much more than literacy and numeracy, it may be argued that it is not the sole responsibility of these subjects to focus only upon a potentially very narrow range of language and number related activities, when they are encountered consistently within other curriculum areas as noted above. It might be expected that English and Mathematics as subject disciplines will continue to reflect very strong impetus for learning language and number related skills, and that within them, certain functional aspects of literacy and numeracy will be addressed in considerable detail.

An illustrative example of this incorporation of literacy within the wider English curriculum may be seen in the following:

"The development of language is central to all learning, whether in 'school subjects' or in any aspect of life. Learning language and learning the possibilities of language through use are central to the English curriculum. Learning how language works, *ie* learning its grammar, in association with its use, is equally important. These things together comprise the field of the English curriculum.". (Northern Territory Education Department, English Curriculum document, 1988).

It might be argued that use of the terms literacy and numeracy as absolute educational objectives is inappropriate in the Northern Territory. Moreover, given that literacy and numeracy can be seen to be much broader terms, with impact upon and from diverse areas of the curriculum, it appears imperative that both are accorded far wider consideration than just within the English and Mathematics subject areas.

The point might also be made that one reason for the apparent decline in students' standards of literacy and numeracy, may be that other subject areas have failed to maintain standards, and that this has consequently encouraged students in the belief that literacy and numeracy are only important in the context of English and Mathematics. That is to say: other subject or curriculum areas have implicitly 'transferred' responsibility for literacy and numeracy to English and Mathematics, and thus no longer accept any major 'responsibility' for teaching and learning within these two fields. This implicit transfer does not suggest culpability, rather that the very weight and intensity of other curriculum areas may now be such that any sustained focus upon numeracy and literacy outside English and Mathematics might be made very difficult.

Chapter 3

System-wide External Examinations and/or Tests Australia and Overseas

Global patterns[1]

Despite moves away from external, system-wide examinations at the conclusion of compulsory education in some areas of Australia, objective forms of assessment remain generally widespread throughout the world.

In dealing with the issue of external, system-wide examinations at the conclusion of compulsory schooling, it must be borne in mind that there are differences between nations and between states in respect of students' ages.

Within Europe and Britain for example, students usually complete their compulsory education in the year during which they turn 16 years of age - a relatively recent development. It is for this reason that British examinations and certification taken at the conclusion of compulsory schooling are referred to as '16-plus'.

In contrast, Australian students generally complete their compulsory schooling in the year during which they turn 15 years of age, a calendar year younger than their European (and many Asian) counterparts.

Therefore, direct comparisons between educational systems may need to take account of variations in the ages of the students, and of the levels at which external, system-wide examinations or tests are set - at the conclusion of three or four years of secondary school in Australia depending upon the location, and at the conclusion of four or five years of secondary school in Europe and Britain.

Such differences do not preclude system-based comparisons, particularly since what is referred to is examinations and certification which mark the 'conclusion' of compulsory education.

The United States of America

In the United States there is no actual provision of external, nationally based, system-wide examinations per se, with objective assessment appearing to be limited, to the utilisation of criterion referenced tests, which are much more widespread than in either Britain or Australia but in many cases generally serve similar purposes.

Stetz F P and Beck M D (1980) in their discussion of contemporary American education write:

"Every American school system has a standardised achievement testing programme involving an average of five test administrations each year.".

Dreher, M J and Singer H (1984) writing in respect of America, report that:

"Standardised achievement tests are the 'most visible and universally used tests in the United States'.".

Irby (1985) reports that 37 out of 50 states utilised objective, system-wide tests devised by Princeton's Educational Testing Service (ETS). Georgia has a total take-up of ETS developed programmes, and three other states - Massachusetts, New York and Wyoming - have requested the ETS to tailor assessment packages for them Irby further reports that most states have programmes that assess the reading and mathematical ability of students in selected elementary and secondary grades. Fifteen states assess the writing ability of pupils. Eleven assess the learning of students in the sciences.

It may also be worthy of note, that admission to universities and other tertiary institutions at under-graduate level in the United States relies principally upon student performance on the Scholastic Aptitude Test (SAT), rather than upon school based assessment or other combination approaches to matriculation.

The British Isles

While there are internal variations in Britain, largely on the basis of regional differences between England, Northern Ireland, Scotland and Wales, the use of widespread examinations at the conclusion of compulsory education is still prevalent.

Student assessment via examinations was previously employed in a range of subjects offered in the English and Welsh General Certificate of Education at Ordinary Level (GCE 'O' Levels) which marked the conclusion of compulsory education for an increasingly larger group of students at 16-plus, and in the General Certificate of Education at Advanced Level (GCE 'A' Levels) which students generally undertook if intending tertiary level studies. Assessment by examinations was also employed in the Certificate of Secondary Education (CSE) - (usually Mode 1) - which was generally offered to students at 16-plus who were not felt to be capable of achieving GCE level studies, or for whom CSE syllabuses were seen to be more applicable or more relevant.

With recent changes (1986/7) to the education system in Britain, both the General Certificate of Education at Ordinary Level, and the Certificate of Secondary Education have been replaced by a unified certificate reflecting syllabuses and courses which attempt to cover a much wider range of student abilities, and which incorporate developments in criterion and domain referenced approaches to assessment. This single qualification is called the General Certificate of Secondary Education (GCSE) and it continues to employ both school assessed and examination based subjects.

Despite the growth of school assessed and non-examined subjects throughout Britain, there have been powerful and persistent moves to provide objective, system-wide assessment of student performance and attainment, not only at the conclusion of compulsory education but at 'fixed' stages during the course of a student's school life. This is reflected in the recent GCSE (as described above) and in the proposed introduction of a national testing programme which will include system-wide testing using grade related criteria at ages seven, eleven, thirteen and sixteen (cf The Task Group on Assessment and Reporting: A digest for schools, 1988).

It appears appropriate to indicate that a number of institutions and examining boards (*eg* Cambridge) in Britain have a history of providing suitable syllabuses and examinations for use in overseas schools and colleges. Typically these have been offered in former British colonies in Africa, Asia and the Pacific, but have also been made available to other developing nations which do not have the expertise or resources to conduct examinations on their own, and because of the international recognition and acceptability which enable students with these qualifications to gain access to further and higher education outside their own countries.

Continental Europe

While it is not possible to generalise for the whole of Europe, there are areas of the continent in which secondary education is based upon a tripartite system which includes comprehensive high schools, technical high schools and secondary grammar schools (often referred to as gymnasia). Students both elect to enter, and are selected to enter, a school considered appropriate for their abilities and aspirations - often this is accomplished using tests designed to indicate students' abilities and aptitudes. It is generally the case that gymnasia are almost entirely 'academic' in their orientation and the main purpose is to prepare students for studies in further and higher education, while technical schools generally prepare students for entry to technical and trade areas of employment.

Examples of this tripartite system occur in Holland and West Germany, with the former using each of the three approaches, although individual schools may offer quite discrete 'strands' on the same site. Regular and systematic testing is common in Dutch and West German schools, and particularly evident in the technical high schools and gymnasia. These tests are conducted on both an internal and external basis and form a primary element in monitoring student progress and attainment.

External examinations at matriculation level (*eg* Arbitur in West Germany, Baccalaurate in France) in many European countries are common and there appears to have been considerably less emphasis placed upon school based assessment as an alternative to system-wide strategies, particularly for what are seen as academically inclined students.

South East and East Asia

Rather than attempt to describe complex patterns of education spanning a number of countries, illustrative examples are used, viz: Singapore, Hong Kong, Indonesia and Japan.

It is accepted that these examples are not directly comparable to Australia, however each is marked by a very strong educational ethos and in Singapore and Hong Kong by a legacy of British colonial rule, including use of the General Certificate of Education at Ordinary and Advanced Level. There is a very strong emphasis upon assessing student performance and attainment in secondary education via external system-wide examinations and tests, both at the conclusion of compulsory schooling and prior to tertiary level studies.

Hong Kong

In Hong Kong, where there is considerable emphasis on relative educational success, the General Certificate of Education (Overseas) at Ordinary and Advanced Level continues to be employed, and it is possible that this will be replaced by an overseas version of the General Certificate of Secondary Education when this is fully developed. Thus, in Hong Kong there is as at least as much reliance upon external system-wide examinations at the conclusion of compulsory education as there is in Britain, with particular stress upon English and Mathematics.

Singapore

In Singapore, the situation is not dissimilar with the continued presence of the General Certificate of Education (Overseas) at Ordinary and Advanced Level which are undertaken by a large number of students, and widespread use of system-wide examinations and tests including English and Mathematics within the general education system. Moreover, within the Singapore education system, students are consistently assessed by means of examinations and tests, and streaming is much more common-place. In effect, this means that examinations and tests within the junior secondary years, form major, integral elements of determining whether a student will proceed to academically, technically or general education oriented senior school programmes.

Indonesia

Education in Indonesia offers a contrasting situation to that evident within the region generally. Compulsory education ends at the completion of primary school - generally around 13 years of age.

During the final stages of primary schooling Indonesian students undertake tests and examinations which may enable them to enter either Government or Non-government secondary schools. It is generally regarded that entry to Government secondary schools is more difficult than entry to Non-government schools, thus requiring substantially better entrance qualifications. Indonesian secondary schooling is divided into three years of lower and three years of upper, with most secondary students completing the lower years only.

Japan

Compulsory education includes the six years of primary schooling and three years of junior secondary schooling. It may be worth noting that on completion of primary school, Japanese students receive a leaving certificate which gives access to the junior secondary area.

At the completion of junior secondary and compulsory education (*ie* at 15 years of age), Japanese students receive a Lower Secondary Certificate. This is based upon a compulsory curriculum and students are graded using a five-point scale in which A=5; B=4; C=3; and D=2, (the minimum pass mark) and F=1 (a failing grade).

Following completion of junior secondary, Japanese students undertake competitive entry for upper secondary education, with a 98 per cent success rate among those who apply. Entry to municipal schools is via achievement tests set by the local authority, while entry to national and private schools is via their own entrance examinations. On completion students receive an Upper Secondary Certificate.

Australian States and Territories[2]

The Australian Capital Territory

The Australian Capital Territory does not employ external, system-wide examinations at either Year 10 or Year 12, and current advice (personal communication to the author 11 February 1988) strongly suggests that there are no moves at all to re-introduce assessment of this type. However there are strong moves to introduce some form of standardised reporting of student attainment at the completion of Year 10.

For students considering entry to tertiary institutions, the Australian Scholastic Aptitude Test (ASAT) is employed in the moderation of school assessed subjects at Year 12 in order to provide a tertiary admission score (TES).

New South Wales

Indications in early 1988 that New South Wales would cease to offer the High School Certificate for Year 10 students, and replace this with a document entitled the Certificate of Secondary Education were substantially altered following a change of government in mid-year. The early intention was to have this new method of reporting available to all students who ceased school after attaining the compulsory leaving age.

New South Wales will now maintain its existing reference tests in English and Mathematics and expand these to include Science in the near future and possibly other subjects at a later date.

It appears noteworthy that for Year 12 results in New South Wales, examined and moderated grades will be shown separately on a student's certificate.

Queensland

Queensland has not employed system-wide, external examinations for a number of years at either Year 10 or Year 12, and advice available (personal communication with the author 11 February 1988) strongly suggests that there is no current or proposed intention to re-introduce this form of assessment though the matter has been re-examined from time to time.

Queensland's appointment of a Chief Inspector within its secondary system may be indicative of wider concerns about the maintenance of educational standards. A system of inspectors at various levels has also been introduced.

Assessment at Year 10, as in other years of secondary education in Queensland, is through school based moderation with no use of reference tests or examinations.

Tertiary admittance scores (TES) at Year 12, are derived in a similar manner to the Australian Capital Territory with moderation of school assessed results by the Australian Scholastic Aptitude Test (ASAT)

South Australia

The South Australian Education Department - (as distinct from the Senior Secondary Assessment Board of South Australia *ie* SSABSA) does not deal with assessment at senior secondary levels and its policy on student performance and attainment at the junior secondary level is being revised. South Australia does not employ system-wide, external examinations at Year 10, and currently appears to be more committed to alternative forms of student assessment including moderation.

It is also interesting to note the very recent (late November 1988) advertisement in the education press for a chief officer to head its new Review Unit which is charged with ensuring educational accountability throughout the State.

(Readers may also wish to note that the role, scope and functions of SSABSA have been the subject of a review, commonly known as the Gilding Report, and that it may be subject to change according to Government policies on education.)

Tasmania

Unlike other Australian states, Tasmania does not have a history of external, system-wide examinations at Year 10. Where such methods of assessment were employed previously, they related to individual schools which were accredited by the Tasmanian Schools Board for periods of three years' duration.

Discussion concerning external, system-wide examinations did take place in Tasmania during the early 1980's, with further discussion of monitoring numeracy and literacy during the mid 1980's. Following close analysis of prevailing conditions, costs and other factors, concerns about the implementation of examinations lost currency and now appears to have abated almost completely.

In contrast, the Tasmanian Schools Board and Department of Education have experience and considerable time and expertise in developing and evaluating objective, criterion referenced tests in a number of curriculum areas, in addition to attempting to develop grade related criteria for the proposed Certificate of Education. With phasing in from 1989, Tasmania intends to utilise a Certificate of Education (TCE) which will reflect students work from Year 9 until the point at which they leave school. Assessment in the new certificate will be based upon a combination of internal and external - which may continue to employ system-wide, external examinations - factors which employ criterion rather than norm referenced approaches to educational goals and objectives.

It is interesting to note that in Tasmania, Grosvenor (1987) reports:

"there was very little demand for the abolition of external examinations .... " (page 5). In this regard Grosvenor (ibid) also reports: "There are still many matters on which final decisions have yet to be made ..... . Important among these are matters of assessment in Years 11 and 12 .... and of detailed procedures for the moderation of internal assessment in all years." (page 9).

It must be noted however, that the preceding discussion relates primarily to senior secondary certification and assessment and not specifically to Year 10.

Victoria

Victoria will be introducing a new Certificate of Secondary Education (VCE) progressively during the period 1989-92. This certificate will mark the satisfactory completion of 12 years of schooling.

The Victorian Curriculum and Assessment Board (VCAB) newsletter for October 1987 states:

Moderation and External Assessment

" .... student achievement in all units of study taken at semesters 3 and 4 level will be assessed by procedures which involve moderation and external assessment.

Beginning in 1988, VCAB will conduct trials of alternative approaches to moderation and external assessment." (page 2)

Although the VCAB document refers specifically to the completion of 12 years of schooling, it may be noteworthy that it specifies that students will be required to undertake:

"four units of the common study of English; four to sixteen units of Arts/Humanities studies; and four to sixteen units of Mathematics/Science/Technology studies.". (page 2)

It is also interesting to note general similarities between the proposals of the Victorian Curriculum and Assessment Board and recent developments in Britain in respect of approaches to moderation and external assessment described as common assessment tasks (CAT's).

"CAT's will be designed to assess differing aspects of student performance. As a component of CAT's, external assessment will include, where appropriate, tests specific to the study." (Assessment for the Victorian Certificate of Assessment, March 1988, page 1)

CAT's are being trialled during 1988 in order to investigate: a range of assessment tasks; procedures which might attend them; their utility and applicability; and their applicability in various settings.

Within the Victorian context, it was reported (Melbourne Herald, 10 February 1988) that:

"The State Board of Education is considering ways to keep a better check on the performance of Victorian schools.".

Of the strategies encompassed in the Minister's discussion paper, two are of particular note in respect of the Northern Territory position vis a vis the potential introduction of external, system-wide assessment. These suggested strategies include:

- "A comprehensive framework of performance indicators, combining data from the assessment and comparison of students' work, and information on costs and outputs."; and

- "An expanded programme of system-wide assessments of student attainment as used in many American States.".

Western Australia

The situation in Western Australia in respect of the lower secondary years has very recently changed. The Secondary Education Authority (SEA) which was previously extensively involved in assessment up to Year 10, has withdrawn from this area and will concentrate its resources in upper secondary activity. In the SEA's annual report (1986-7) it indicates that moderation for courses at lower secondary will be devolved to the Ministry of Education and schools, and that the SEA will liaise with schools in order to: "support the introduction of a standards-referenced approach to assessment and grading." (page 11). The SEA report also indicates that comparability tests for lower secondary education will be abandoned for 1987 and beyond.

However, it is equally clear that in the upper secondary years, the the SEA will be strongly pursuing system-wide, objective testing utilising item banks created for this purpose in a number of subjects; and that Western Australia intends to maintain tertiary admission level examinations on a system-wide basis. In 1986, the tertiary entrance procedure was based upon examinations, moderated subject assessments and scores on the Scholastic Aptitude Test (SAT).

Chapter 3: End Notes

1 and 2 The information contained in this chapter was accurate at the time of going to press, however it is anticipated that changes in education policy and in approaches to assessment may lead to a reduction in the utility of the material provided. Readers are therefore advised to check up to date information which is produced from time to time by State, Territory and other authorities.

Chapter 4

Cases Pro and Anti System-wide External Examinations and Tests

Precursors to the debate

Arguments for and against external examinations and tests, and about examinations and tests per se have raged for a considerable period of time. The only general concurrence amongst the protagonists and antagonists in this argument may be in agreeing to differ.

Moreover, the debate about student assessment, measurement of performance and attainment and about the utility of certain educational practices has been clouded by unsubstantiated opinion, unsupported assertion, political expediency and elements of personal and professional vitriol; with arguments about the relative advantages and disadvantages of differing forms of student assessment obscured by point-scoring and stance-taking. Indeed, arguments are often based upon philosophical or ideological stances rather than educational criteria, with apparent influences from adherents of the humanist and the opposite approaches to society. Thus, the 'general public', and indeed many educators, have formed views on various models and methods of assessment which are based more on untested assertion than upon documented, rational evidence.

It may also be true, that proponents of alternative non-examination based assessment systems, have done themselves a considerable disservice in failing to address the critical question of informing their audiences and stake-holders about the validity of their claims.

In contrast during periods of system-wide, external testing and examination, the public - comprising students, parents, employers and others - were generally made aware of what the examinations sought to do, how they functioned and what standards students had attained. Examinations and tests conducted were 'validated' by ad hoc perceptions of student worth, impressions of literacy and numeracy prevalent at the time, and by the simple expedient that they were so long-standing and widespread that nearly everyone who had completed primary or secondary education either in Australia, or overseas had some experience of them.

This validation must be placed in a context of examination completion rates. For example,in Britain prior to 1950 no more than 10 per cent of students actually completed the School Certificate, and the General Certificate of Education Ordinary level (GCE '0' level) was originally designed to meet the needs of 20 per cent of the secondary school population.

In addition, the universities, which were regarded as the apex of educational achievement, all employed examinations as their principal means of assessing student progress. That these were not genuinely external, was of little moment. The presence of annual or more frequent examinations in the university system added a perceived legitimacy to their use in wider society. The same cannot be said for alternative forms of measuring student performance and attainment, particularly moderated school assessments in schools and tertiary institutions.

In general terms, the models and methods employed have been singularly lacking in their 'public' presentation, and there appear to have been few detailed attempts to explain any advantages which accrue through school-based assessment, to parents and employers who are much more familiar with traditional examination systems. This lack of 'publicity' (read *public relations*) has been compounded by the relative recency of non-examination based approaches to assessment, which deprives them of the legitimation of time and familiarity brought about by longevity.

In short, whilst today's claims about poor student attainment and low standards of literacy and numeracy, may be as accurate or as inaccurate as they were in the past, education systems which rely upon non-examination based approaches to student assessment seem less well able to defend themselves, and thus appear to lack credibility when responding to allegations of poor relative performance.

In addressing questions which pertain to system-wide examinations and/or tests in the following sections, it is suggested that the broader role of assessment in schools and other institutions of learning be borne in mind. In particular, that student assessment is an essential and vital element of the processes of teaching and learning, and that it should not function as a mere adjunct which may be useful or not.

If the central role of student assessment is accepted within teaching and learning situations, then the debate should not be so much about whether or not to assess, but more about the mechanisms which are employed to do so. As far as assessment is concerned, the key to the perceived validity and wider recognition of any formal examination, test or other method, lies in the questioning and marking techniques used.

Arguments may also be advanced for the separation of formative and summative approaches to assessment and reporting. That is to say: it may be argued that there is a need or demand for a formal 'report' or assessment at the conclusion of a period of study - the completion of compulsory education for example - which is quite distinct from the processes employed to assess and report and guide teaching during that period of study.

System-wide Examinations and Tests: a Case Against

This section provides a summary of the arguments which have been ranged against either the continuation or introduction of external, system-wide examinations and tests, particularly at the conclusion of compulsory education.

In much of the recent, national and international discussion vis a vis examinations and testing, there appears to be an assumption that present education, and present educational standards, are somehow worse than in earlier periods and in some other countries and states, and that the introduction of examinations or tests will in some way redress this apparent imbalance.

On the other hand it is argued that, no matter how well designed, examinations and tests are, they may not be effective or efficient measures of the curriculum, or appropriate to other than small groups of students at any one time. Examinations and tests may 'measure' or 'test' only part of a school or system curriculum at any one time, given limitations upon time and other logistic concerns. Moreover, that an examination or test, may not in fact provide an effective means of assessing the knowledge a student possesses and may therefore present a limited, and possibly incorrect picture of performance and attainment.

This point of view is advanced in the submission made by a number of respondents including the Northern Territory Teachers Federation (NTTF) in responses[1] to Towards the 90s: Volume 1. The Northern Territory Teachers Federation document states: "An examination can at best, only sample a narrow range of skills and knowledge acquired during an exhaustive three year programme.". It is this suggested narrowness of focus which gives rise to concerns about what, and how effectively, an examination measures; and about the extent to which an examination based assessment system may limit the curriculum through encouraging teaching solely for the examination.

Dockerell (1984) addresses the competing claims of external, syllabus based assessment, and internal assessment. He states that: "External syllabus based assessments have been criticised on the grounds that they impose an intended curriculum and thus constrain teacher initiative by imposing a unitary curriculum on diverse local circumstances.".

Dockerell also states: "One of the advantages asserted for school based assessments on the other hand, is that they relate to the implemented curriculum, that is, to what the teachers actually teach in their particular circumstances and not to the notional intended curriculum".

This argument is extended when it is claimed that tests/examinations may bring about a lack of balance and relevance in the curriculum and in the ways in which knowledge and learning are measured and valued.

It is further argued in the Northern Territory Teachers Federation submission (with a similar tenor in the comments also made in other submissions) that:

"The prospect that such a measure (*an examination?*) should be accorded equal status with the individual school's carefully moderated continuous assessment programme is lacking in balance.".

Although the Northern Territory Teachers Federation is clearly referring here to the balance between forms of assessment, the notion can readily be extended to possible distortion of the curriculum in order to concentrate upon those items known to be, or felt likely to be the subject of examination, at the possible expense of other equally meritorious items which may not be assessed in this way.

It may also be the case, that items which are assessed by external examinations or tests are accorded higher status or greater significance than those which are assessed in other ways. This may arise from public and educational perceptions of 'worth' and 'value', but equally from an underlying notion that if an item is tested, it is important, if it is not tested then its importance is diminished.

One corollary argued in relation to the above is that examinations and/or tests act to reduce or prevent sound learning and teaching practices by encouraging in students rote learning techniques, and over-concentration on prepared or 'stock' answers to likely questions and topics.

In addition to concerns about possible imbalances caused as a direct result of examinations and tests vis a vis other forms of student assessment, it has been claimed that there a number of other factors which militate against the introduction of external, system-wide examinations and/or tests.

First

The introduction of system-wide, external examinations or tests is counter to trends evident in other education systems at Year 10 level - particularly those within Australia - and to recommendations made some time ago by the Northern Territory Board of Studies. This assertion is certainly sustained by the evidence available from other Australian States and Territories, which, with the notable exception of New South Wales, either abandoned Year 10 external, system-wide examinations some years ago, or, as in the case of Tasmania, never employed this particular mode of assessment.

In this sense then, the Northern Territory would represent a something of a unique example in offering examinations and/or tests at Year 10, - the conclusion of compulsory education - as well at the end of Year 12, and as an integral element of tertiary entrance procedures. It would also be only the second education system utilising examinations or tests for the relatively small percentage of students who actually cease their full time education at Year 10, and who do not proceed to Year 11 and/or 12. It is possible however, to view Year 10 assessment not in the form of a 'leaving examination' but as a qualification for completing a distinct phase of schooling which enables students to proceed further.

Second

It is argued that the introduction of examinations and/or tests may encourage a competitive, rather than co-operative, classroom environment, in which the advantages of shared learning and whole-group based approaches to learning are diminished.

Third

Examinations and/or tests may bring about, and institutionalise failure, rather than produce an overt focus on development of successful academic, social and personal skills for all students.

Examinations bring about a form of ranking which is either imposed upon students or which is perceived as students compare their work with that of their peers. Examinations and tests as relatively simple methods of assessment based on ranking may also act as a stimulus or impetus for a self-fulfilling prophecy in which certain individuals or groups are expected to perform badly and then do so, a 'fact' which acts to confirm teacher and societal expectations.

The pressure of examinations against co-operative learning implies that the general focus of the classroom will be the teacher who may be seen as the owner of knowledge, with consequential diminution of the role students as peers may play in facilitating learning.

Fourth

Examinations and/or tests may be one factor which frustrates curriculum reform and development throughout the school. This will be due to concentration on centralised syllabuses or curricula, which may be provided in the form of prescriptive statements which limit the extent of school based adaptation.

Moreover, it is possible to contend that the Education Department of the Northern Territory and the Board of Studies have actively encouraged schools in the development of courses and programmes in Years 10, 11 and 12 which seek to meet student needs and demands, and that examinations and tests will act in a way which is directly counter to this long established approach.

It its submission, the Northern Territory Teachers Federation (together with a number of other respondents) argued that the introduction of external examinations would relegate sound educational practice and courses in other curriculum areas to second rate status.

In addition, it can also be argued that the introduction of proposed examinations and/or tests may disrupt the life of the school community and may damage the quality of education provided. This argument takes on more immediate impact when it is considered that the stated focus of external examinations and/or tests is upon only two curriculum areas viz: English and Mathematics. One possible consequence of this position may be a parallel to that noted earlier, that is to say: English and Mathematics assume greater significance and critical importance in the view of the general public, while other areas of the curriculum are reduced in perceived status.

Indeed, it may be difficult to sustain a rational approach for examinations in English and Mathematics and not for other curriculum areas which are currently prescribed as 'essential' - Social and Cultural Education; Business Education; Science; Technical Studies; Home Economics; the Arts; and Languages other than English.

Fifth

It has been argued that system-wide, external examinations and/or tests may function as a selection and channeling mechanism, through the use of which structural and social inequalities are maintained. In essence, examinations and/or tests act to selectively deny student access to particular senior secondary courses, and beyond them to further and higher education. Examinations and tests as an integral element of education systems, may impinge upon the life chances of groups including women and girls, and members of ethnic or socio-cultural minorities.

The Northern Territory Teachers Federation and other respondents, in their submissions in respect of Towards the 90s: Volume 1, have argued that examinations produce severely biased outcomes on a social class, gender and ethnic basis, and support the same bias in relation to tertiary selection and employment prospects. Moreover, there is research evidence to support this view (cf Jackson and Marsden 1968, Husen 1972 and 1975, Robbins 1973, Ntuk-Idem 1978, Reid 1977 and 1981, Lazar et al 1977 and 1978, Mortimore and Blackstone 1982)

If the preceding argument is accurate, then the adverse impact of examinations upon large sections of the community is both unfortunate and unwarranted. However, it could be argued that it is achievement per se which should be assessed, not potential. This paper is not seen as the context in which to pursue the essential question of whether education systems should even attempt to measure potential or to redress social and other disadvantage.

Sixth

There does not appear to be strong and prevailing evidence which sustains the notion that examinations and/or tests are a critical way in which educational standards might be raised. Examinations and/or tests measure the position as it is, rather than as it might be. The notion that the presence of examinations and/or tests by themselves will act to raise educational standards does not appear to have a sound basis. It is the nature and use of examinations which will determine this result.

It would seem more likely that standards will improve consequent upon a number of factors, examples of which may be: school facilities and resources - both material and human; strong relationships between school and home/community; a highly motivated and professional teaching body which is stable and well supported; and other effective school factors and approaches.

One possible extension of the preceding discussion suggests that educational systems which seek to raise educational standards by requiring students to pass minimum competency tests to advance or to graduate, may in fact inculcate a speedy way to lose the disadvantaged youth which society has sought to retain within the school system. Alternatively, the use of tests and/or examinations may assist in identifying those who require 'special' treatment.

In effect, this argument holds that by valuing examination success above all else, or by placing such examinations or tests in a pre-eminent position within an education system, a direct result may be the alienation of those for whom they have little relevance - chiefly those students who will fail to do well.

If the above is accepted, then it may appear that the phenomenon of examinations, testing, grading and scaling bring about unfair and inaccurate assumptions of individual worth.

This happens because they tend to emphasise an isolated factor, or set of factors generally established as 'academic worth' at the expense of other important social and behavioural developments. That is to say: examinations and tests may undermine the efficacy of teaching and the development of socially desirable characteristics in students by their very emphasis on other teaching and learning approaches.

Seventh

It may be argued that examinations and tests deny the reality that students and other individuals, make progress at different rates and in different subjects.

The traditional approach of matched age and grade assessment in education has been challenged in recent years, and it is now more widely accepted that there may be significant differences in the rates at which people learn, levels to which they might aspire in different subjects or curriculum areas, and in human attributes and aptitudes for different aspects of learning.

Recent moves towards systems of vertical time-tabling and/or unitary curricula in the Northern Territory and in other States, appear to have as one element of their rationale these differences in learning patterns and strategies. Proponents of alternative curriculum and school management approaches have argued that non-test or examination oriented settings allow for individual variation and the matching of a given subject, curriculum area or syllabus to personal and individual need.

Moreover, within contexts of this type, it is argued that examination achievement may not bear any significant association with overall student performance. The rationale underlying this proposition is that schools have, and must continue to maintain, more than a singular and potentially very narrow approach to education which is 'measured' by examination success.

Schools are seen as intrinsically valuable in that they offer education in its broadest sense through a combination of effective and affective learning, human growth and development which is both physical and psychological, and are responsible for preparing young people to take a place in the wider community.

Support for the preceding may be found in Bresser (1987) who, in writing on national testing states:

"Learning takes place at the individual's own pace which must be monitored and recorded, and recognition given for achievement. It is impossible to expect that each child in any age cohort will be at the same stage of learning and development.".

As was noted earlier, examinations and tests may carry with them some concomitant responsibility for the promotion of a potentially more restrictive curriculum and for the promotion of a narrow view of the goals and objectives of education. It might be argued that examining only a limited number of skills or aspects of the curriculum cannot and does not indicate an individual's experience and abilities in the whole educational spectrum.

Eighth

There appear to be quite strong arguments in favour of the point of view that examinations and tests may be unreliable and invalid indicators of student success in career or later education.

If this suggestion is accepted, then the potential value of examinations as predictors of educational or vocational success assumes a much more limited perspective, with the consequent need to incorporate other more reliable means of such prediction.

Ninth

One of the critical dilemmas which attends external, system-wide assessment concerns the utility and application of results. That is to say: it appears quite unreasonable to make school by school, or similar comparisons of the results of student attainment without allowing for other variables which may influence such outcomes.

While this problem does not reflect upon assessment per se, the consequences of using results in a generally superficial manner can provide a distorted, unbalanced and inaccurate portrayal of student performance and attainment. Studies undertaken under the auspices of the former Inner London Education Authority (ILEA) and those noted earlier (Gray, Jesson and Jones ibid, Goldstein ibid, and Fogelman ibid) suggest very clearly that there is a need to incorporate background factors such as home circumstances, socio-economic status and ethnicity in any analysis of student performance and attainment. By extension, such studies also suggest the need to take account of the potential uses of assessment information prior to implementation.

System-wide Examinations and Tests: a Case For

It is difficult to argue that the factors of time, history or tradition are themselves adequate defences for system-wide, external examinations.

Further, their widespread use geographically does not necessarily offer sufficient reason for educators and others seeking their introduction or re-introduction in local settings. However, cognisance must be taken of the utilisation of examinations, tests and objective measurement in so many settings, and this leads to a central question:

'Why, if examinations and objective measures of this type are so educationally unsound, does their use persist in so many countries of such varied backgrounds?'.

It is necessary to turn to the evidence and rationales which underlie the use of examinations and objective measures of student performance in the many settings in which they are found, rather than rely upon untried assertion and assumption.

One of the key arguments which underpins the utility and application of both examinations and other forms of assessment, is their use in monitoring student performance.

It was argued in Towards the 90s: Volume 1, and it appears beyond dispute, that:

"The Department of Education has a responsibility not only to continually to strive to raise the standards of literacy and numeracy of the Northern Territory school population, but also to increase public confidence in the results and qualifications achieved by students in schools' programmes.". (page 9)

System-wide, external examinations and/or tests are aguably one means by which educational performance in key curriculum areas may be assessed. Such examinations and/or tests may focus upon particular curriculum goals and objectives, and can be used to provide indices of student performance and attainment in areas which are considered critical. Moreover, these indices can be presented in a manner easily comprehended by educators and the public alike.

That is to say: while an examination in the Northern Territory may not seek to assess student performance in the 'whole' curriculum and over the 'whole' initial three years of secondary education, it may focus upon those aspects which are identified as Core Objectives and which are incorporated within Board Approved Curriculum.

System-wide examinations and tests may also be employed to assess student performance and attainment in those areas of literacy and numeracy which have given rise to greatest public and educational concern, and may therefore provide satisfactory evidence acceptable to the wider community of how well, or how poorly, Northern Territory students are achieving. Examples of these critical areas include: reading; writing; listening and speaking; and essential arithmetical skills.

In other words, if the power of examinations and/or tests to focus efforts upon particular subject or curriculum areas is acknowledged, this power may then be used productively to ensure greater attention is accorded to those features which are considered to be most important.

In essence then, there appears to be a need to identify and assess what knowledge and skills a student should attain at school. Bourke (1982) in an article dealing with testing rather than examinations, argues that:

"It is in these areas of fundamental learning where it is possible and justifiable to implement a survey testing programme. The fact that all schools are teaching the content and skills tested is indicative of some consensus on the importance accorded to these

areas. It seems reasonable that these are also the areas of greatest community concern about the outcomes of schooling. Survey testing in these areas would not only recognise their importance but would also help to re-assure the community that everything was being done to ensure that the highest possible proportion of students would become literate and numerate.". (page 79)

While Bourke's article refers principally to testing, the comments appear at least partially applicable to external examinations.

Support for the use of system-wide examinations and testing also arises from diverse educational settings. O'Loughlin, Bardsley and Bird (1984), report that in the United States:

"Student achievement in external examinations has long been used by teachers, administrators, parents and others to evaluate the academic performance of their own school. Externally set and assessed examinations are seen as valid and reliable indicators of the quality of a school's teaching programmes.".(page 1)

This argument amplifies the utility of examinations beyond providing public information about student performance and attainment in specific curriculum areas, to an evaluative role in which a school and the community of which it is a part can obtain reliable indications of the success of its teaching and learning programmes.

It is very the isolation of a teacher in a classroom which precludes all but a few bases upon which to judge how well they, or their students are performing.

By extension, system-wide external examinations offer two critical reference points:

firstly a summative assessment of student performance and attainment in key curriculum or subject areas; and

secondly a form of diagnostic assessment for individual schools and the whole educational system which may indicate relative strengths or weaknesses in teaching and learning.

The dual roles may be further added to by the inclusion of the processes of educational legitimation and certification which are critical in the pursuit of further and higher education and in the take-up of employment.

This amplification is addressed by Adelman (1984) who writes that:

"The use of tests to sort and screen is legitimate for professional certification and licensing - and indeed, for any operation where selection is necessary. But assessment has even greater potential as a tool for clarifying expectations and for increasing student involvement when it is used to measure improvements in performance."

The second element of Adelman's statement suggests that far from driving students away from the processes of teaching and learning, it is possible that assessment may enhance student involvement in these processes, particularly when it enables students and teachers to measure improvements in performance.

It may be further argued that Australian society, in addition to claims made in respect of its egalitarianism, is essentially a meritocracy in which access to further and higher education and to all occupations does not generally result from caste membership, social class or other imposed structures, but from individual aspiration and effort. In this respect Australia shares with other Western nations the fundamental ideal that selection for employment, education or training should be by combination of effort, aptitude and ability, all of which are demonstrable in a variety of ways.

A parallel position vis a vis meritocracies, may be observed in the United States, a nation which has done a great deal to democratise its own education systems, and in which the use of testing is almost universal.

It is in the context of both a meritocracy and state/system-wide testing that Irby (ibid) has written:

"A society in which opportunity is open to anyone with high ability who is willing to work hard is a long standing goal -- and ideal -- of the US education system. The educational system, given its capacity for identifying and training those with exceptional talent, is crucial to the functioning of this kind of meritocracy. In this context, the fair and equitable assessment of students is one of the foundations upon which the entire enterprise rests.". (page 1)

In the United States, as in other western European countries, the use of examinations and tests as measures of student performance and attainment remains not only widespread, but an integral element of the whole educational process.

An illustration of the extent to which tests and examinations influence educational 'thinking' in the United States, particularly among teachers and practitioners may be seen in the following examples. Stetz and Beck (1980) report that of a national survey of 3,500 teachers in the United States only 16 per cent agreed that standardised testing should be abolished; and the National Education Association (1980) found that: 90 per cent of teachers responding said that tests were useful as long as they were not used in isolation, and 84 per cent said that they could not teach effectively without them.

The testing programme referred to is carried out by the National Education Association which conducts relatively frequent and systematic nation-wide tests throughout the United States. National educational patterns are inspected by the NAEP, which samples 100,000 students throughout the country - it measures reading ability every two years, mathematics and writing every four, and other subjects less frequently. The NAEP programme is designed not only to measure student performance per se, but to enable education departments and schools to assess achievements on a national and state-wide level and to spot trends in curriculum areas, in geographic regions, for certain sub-populations, and in specific knowledge and skill areas.

This overt reliance upon testing which is noted in the United States, is reinforced by the reports of Government sponsored commissions of enquiry and by other organisations with considerable experience in the design and administration of testing programmes.

Illustrative examples of this support may be seen in the report produced by the National Commission of Excellence in Education (1983) which is entitled A Nation At Risk. This study warned of a "rising tide of mediocrity" and recommended that "Standardised tests of achievement ... should be administered at major transition points from one level of schooling to another and particularly from high school to college or work.".

It is these transition points from primary to secondary, and from lower to senior years which appear to be causing the gravest concerns in the Australian, American and British educational contexts, and it is the transition from junior secondary to senior secondary which is a point at issue within the Northern Territory.

A Nation At Risk also recommends that such tests could be used to certify student credentials, identify the need for remediation, and identify opportunities for accelerated work.

A further example of support for testing may be seen in statements made by William Turnbull, formerly President of the prestigious Educational Testing Service (ETS) based at Princeton University. Turnbull argues that: "the revolution of falling expectations" causes even the best students to work less hard and learn less, because less than the maximum performance was sufficient to earn top grades, and that: "a curriculum, if it is to work, must be sustained by an appropriate and rigorous system of student assessment.".

Turnbull suggests that in the USA and elsewhere, 'declining' educational standards - particularly in university entrance level (SAT) scores have been attributed to a number of factors, or remarkable similarity to those noted within the Australian context.

These include:

the simplification of language during school;

a broader range of social classes and other groups taking the tests rather than the former academic elites - which is equivalent to dilution of the groups who may have once taken such tests;

the proliferation of elective courses at school enabling students to avoid tough academic courses;

the influence of television including extensive viewing times; and

poorly prepared students being retained in schools but still seeking entrance to further and higher education.

Statements by Turnbull and those noted in A Nation At Risk, are strongly supported by other American studies which include: the Carnegie Foundation High School: A Report on Secondary Education in America.; the Task Force on Economic Growth Action for Excellence; the Commission on Pre-college Education in Mathematics, Science and Technology Educating Americans for the 21st Century; and the Twentieth Century Fund Making the Grade.

Within the United States these reports were generally greeted with almost universal approval and endorsement from the general public and from a substantial portion of the nation's teachers.

In Britain, the whole issue of examinations and testing has re-emerged as a significant element of contemporary education debate. Earlier within this paper, note was made of objections to the use of examinations by authors including Howson (ibid) and Murphy (ibid).

While examinations in Britain have their detractors, they also have their defenders, of whom Syred (1987) is one. In an article which appeared in The Times Educational Supplement (13 November 1987), a systematic rebuttal of the criticisms of examinations was made.

Opponents of examinations and tests have claimed that external, system-wide assessment is likely to be disruptive for school administration and educational programmes. However, a countervailing argument holds that this assumption of disruption must be placed in a context where the quality of education is so uniformly good that disrupting it would be unjustifiable.

It does not seem possible to sustain an argument that any educational system, including that in the Northern Territory, is so ideal as to warrant no external disruption whatsoever, and particularly one which seeks to measure just how well the system is performing.

The assertion that examinations and tests will not raise educational standards, but rather measure what already exists, appears to be a naive and overly-simplistic approach. That is to say: the very basis for using a criterion referenced system to determine whether a particular level has been reached, strongly suggests that it should be employed totally rather than selectively; if objective standards are set, and these are system-wide, then it may be argued that objective, system-wide measures appropriate to the assessment of those standards, should be employed as a means of determining which of the population is achieving or not achieving.

Earlier (cf Howson, Murphy, Northern Territory Teachers Federation and other arguments), it was asserted that in addition to their inherent disruption, examinations would almost certainly bring about a narrowing of both the curriculum and of teaching per se. In essence there are a number of counters to this approach, one of which - teacher professionalism - seems to have been almost entirely disregarded.

If teachers are genuinely professional, it is inconceivable that they would concentrate solely upon those subject or curriculum areas in which an examination scheduled, at the expense of all other, equally valuable components of the curriculum. It is equally inconceivable that responsible teachers would effectively doom their students to failure by not making every effort to ensure that the students performed as well as possible in any external examination.

On the other hand, it is recognised that in some settings, teachers have acted to exclude from their examination classes, those students who they felt would do less well, and thus detract from the teacher's success record. However, both past experience in Australia and current tertiary entrance level studies, and experience overseas, neither supports the narrowing of teaching proposition, nor any general lack of professionalism which teachers bring to the classroom.

That is: the 'evidence' used to counter examinations or tests selected from the past, is often based upon faulty examination systems, or upon particular de-contextualised remarks and comments which are then generalised. Further, this supposition is largely confounded by the principle that at best, all teachers can do is 'second guess' examination questions. For teachers to risk leaving unexplored large or critical areas of a curriculum is tantamount to courting disaster for their students, and possibly themselves.

It is also appropriate to point out that within some systems teachers who gain high levels of examination success are generally those who are awarded high professional status.

Within this discussion, it must be seen as an untenable proposition that teachers who do their utmost to prepare students and ensure they do as well as possible in important external examinations, are in some way showing a general lack of professionalism.

As was noted earlier, much of the anti-examinations rhetoric focuses upon the notion that teaching and the curriculum will be narrowed, with some subjects gaining precedence over others.

One of the essential principles of education per se and of examinations and/or tests as an integral element of assessment, is that there should be appropriate relevance and balance in teaching and learning. There is nothing in this principle which suggests primacy for some subjects and subordinate status for others. Therefore, arguments about the status of examinable subjects to the detriment of non-examinable subjects may be seen as largely baseless, unless examinations are confined to a narrow range of subjects.

In terms of balance, there is no implicit notion of equality, more an acceptance of equity which accepts that students' needs vary over time, and with individual circumstance, thus requiring a dynamic rather than static approach to education. What may constitute a major element of a student's educational life in early primary, may be substantially diminished by the late secondary years. Moreover, it is widely accepted that certain classes or types of subjects should receive preferential placement, or have the 'balance' tipped in their favour.

This is recognised within the Northern Territory (and in many other locations *eg* the Task Group on Assessment and Testing in Britain), where a defined group of subjects and their curriculum are mandatory from early childhood to Year 10, while others are treated as optional or electives which may be taught according to demand or availability. But even within this group of subjects most would probably agree some deserve more emphasis than others.

One of the key areas in which system-wide examinations and/or tests have been attacked is in their acceptability as assessing tools. This usually happens in a context in which moderation of school assessments and other forms or continuous assessment are equated with 'goodness', and examinations and other objective measures with 'badness'

In no logical, rational or educational terms can such over-simplification of 'goodness' versus 'badness' stand scrutiny or inspection. It is possible to have both good and bad examinations and tests - that is to say: those which are valid and reliable and those which are not; similarly it is possible to have good and bad alternative forms of assessment and moderation to which the same constraints must always apply.

Further, it cannot be argued that because examinations are introduced or re-introduced, and are conducted on an annual or bi-annual or other basis, that other forms of assessment must of necessity cease to operate. There is simply no immutable law which demands an either/or situation. That is: if a system has examinations this should not preclude alternative assessment approaches.

An ideal approach to assessing student performance and attainment might be one which is eclectic and takes up the best options from any and all available models, without excluding any other on tenuous political, philosophical or educational grounds which may lack foundation. This lack of logic often found in relation to examinations is surprising given their potential utility in diagnostic and formative approaches to student learning and attainment. It is difficult to conceive of clearer information for teachers, than that offered by a systematic, well designed, valid and reliable test or examination. But this aspect of assessment is almost entirely disregarded in the arguments of those opposing system-wide examinations and/or tests.

A factor closely allied to the 'informing' rationale of examinations and tests in respect of teaching and learning, is the right of educational stakeholders - students, parents, teachers, employers and others - to know just what is happening in schools and colleges. Students, parents and others should have the right to comprehend the relative position of an individual or school against a system-wide or national norm. It appears quite unjustifiable, that parents and others are not informed unequivocally and straightforwardly, about system-wide attainment levels, against which comparisons might be made.

It is appropriate at this point to indicate that secondary schools in England and Wales are now required to publish their students' performance in the annual '0' level and 'A' level assessment. This was made mandatory under the provisions of the Education Acts 1980 and 1983. Recently published papers (Gray, Jesson and Jones 1984, Goldstein 1984, and Fogelman 1984) in respect of student performance and attainment at CSE/'O' levels *ie* 16 plus in Britain, show quite conclusively that unless background variables *eg* social class, socio-economic status or non-English speaking home life, associated with the intake populations of schools are taken into account, the publication of results can be very misleading.

To fail to provide information about student performance and attainment in a way which is able to be comprehended by the public at large, smacks of introspective professional isolationism - an attitude that may be at the bottom of our schools' 'public relations' problems.

Examinations and/or tests often appear to be opposed on the grounds that they somehow deny the reality that individuals progress at different rates and in different subjects. There are a number of strong rebuttals of this argument, examples of which are illustrated below.

• It is accepted that individuals do learn and achieve at different rates and in different subjects, however this must be contrasted against the reasonable demand that, by the completion of their compulsory education, all students have achieved certain, pre-determined levels of performance and attainment in specified subjects.

• This position is not dissimilar to that applying in tertiary admission, which requires students to undertake entrance tests, examinations or other forms of assessment (for example America's SAT, the Australian Scholastic Aptitude Test, and Britain's 'A' levels), age, gender and other factors notwithstanding.

• The objection vis a vis differential learning is much reduced when it is considered that it relies upon a notion of adding results from different subjects into some 'total' mark or grade which denotes the conclusion of a particular period of schooling. With the exception of ASAT/SAT type tests in which some 'compounding' does occur, this would be a very unusual and difficult procedure.

A response to the notion that it is unfair to establish 'targets' for students and teachers to aim at, or to place barriers in the path of an individual's progress, must be the question *'why?'*, and why should we not seek to extend all students? It is widely accepted that individuals have differing talents and abilities, together with differing aspirations and objectives. To suggest that these can be 'engineered' out by removing any barriers to entry is both naive and potentially very dangerous. Most trades and professions decree minimum entry points, related to, or based upon, possession of educational criteria, and many upon examination or entry tests - hairdressing, catering, motor mechanics medicine, law and engineering are examples. Removal of these 'barriers' may not of itself open the trade or career to everyone, but it may act to reduce the intellectual or associated quality of those seeking entry to the point where some practitioners can damage human life or property.

It may also be argued that social engineering cannot effectively replace the essentially meritocratic nature of western democracies, and the concept of universal mediocrity is unpalatable to many people. It seems equally unpalatable to argue that because individuals progress at different rates, that this should somehow lead to the proposition that they should not be encouraged to perform and achieve at their maximum.

There is something profoundly disturbing about suggestions which appear to promote 'averageness', and which do not entertain the principle of ensuring optimum education and related achievements on the part of young people. This seems acceptable in strategies to help disadvantaged or handicapped students, but not those who are gifted and talented.

In educational, rather than rhetorical terms, it has been argued that examinations and/or tests may be both brief and unrepresentative in respect of students' learning and attainment. This argument holds that examinations and/or tests cannot 'measure' a subject or curriculum, and that they may present a false or distorted picture of student performance and attainment as they are only concerned with a small part of intended learning.

The counter argument would hold that an examination and/or test which is both valid and reliable will provide an accurate and meaningful measure of student performance and attainment in those areas of the curriculum which is competent to measure. That is to say: the domains of an examination or test are only those to which it is applied. There are dangers in over-generalising. An examination and/or test should be closely related to the objectives and content of a subject or curriculum and should measure attainments related to these. Student results should only be interpreted in terms of what the examination or test sought to measure.

An examination which sought to measure factors unrelated to the curriculum would probably lack both reliability and validity, achieving any significant result purely on the basis of chance. It would also be expected that the correlation between the results of an examination and/or test based upon curriculum content and objectives, and other assessment measures would be strong and highly significant, and that this would prevail even if a limited number of curriculum aspects were employed as a base.

Effective examinations and tests are akin to other effective assessment measures in that they seek to show both what is happening in student performance and attainment. The notion that they must somehow 'test' only a narrow or limited range of factors, and that this is therefore inappropriate does not stand counter argument.

Most complex and varied phenomenon can be reduced to a number of possibly discrete and/or inter-related elements (cf Templin's model for tests of measured intelligence, and those employed by Vernon, Heim et al). It can be shown that strong and highly significant correlations (or other measures of association) exist between pairs or clusters which enables tests of certain elements to act as 'representatives' for a much wider subject or curriculum. That is to say: the results of tests of various components of a compulsory mathematics syllabus would be expected to produce strong and highly significant correlations; they would also be expected to inter-relate strongly with other assessment measures; and most importantly, with a notional or measured ideal of student performance in the 'whole' subject.

A further way of capitalising upon the reduction in the number of test or examination items, is to employ fewer of them, but to give weight those which are considered important or critical. In this way, the 'central' or essential curriculum might be assessed, with lesser weighting to peripheral or subordinate material.

One criticism of 'traditional' examination patterns is that they bring about a certain degree of 'coarseness' in grading, usually as a result of a relatively low number of categories. This coarseness is held to be disadvantageous to certain groups or individuals. It must be pointed out that this situation is not confined to examinations, but applies equally to other forms of assessment which employ limited scales - for example grades A to E, or other five-point models.

While low-number scales may compress scores to 'fit' distributions of this type, one simpler solution is to extend the scales from five-point to ten- or twenty-point. This last strategy has been in use with the Senior Secondary Assessment Board of South Australia (SSABSA) for some time. Extended scales of this type reduce the possible compression of marks to 'fit', and are therefore capable of providing greater expressions of discrimination between student performance and attainment.

Opponents of examinations have asserted that reports of assessment of this type are disadvantageous to: members of ethnic minorities and non-mainstream socio-cultural groups; low socio-economic status or low social class students; and female students. Presumably, opponents of examinations and/or tests would argue that those most disadvantaged by examinations will be females from an ethnic or non-mainstream background, who are also from low socio-economic status homes.

It is here that arguments vis a vis examinations and/or tests appear to have become confused with those surrounding tests of measured intelligence, and with a number of unsustainable claims in respect of particular groups within the wider community made solely in respect of such tests of measured intelligence *eg* the whole 'culture fairness debate surrounding intelligence testing.

Evidence from Britain (Douglas, 1964), strongly supports the use of objective assessment in respect of the equitable selection of candidates from differing social class backgrounds. Douglas reports that when objective examination measures were employed, the proportion of young people selected for grammar school entry (11 Plus) from working class homes, was very nearly identical to the proportion identified as eligible by tests of measured intelligence. In contrast, when subjective measures were employed for selection at 11 Plus - including teacher identification - the relative proportion of working class background students decreased, and the relative proportion of middle class students increased.

That is to say: it was when students were predominantly identified by more objective measures of intellectual and academic performance - including tests - that working class representation in selective entry grammar schools very nearly matched the numbers indicated using tests of measured intelligence. Nevertheless, an element of bias may still remain since it can be shown that both measured intelligence and academic attainment are positively correlated with measures of social class, status or home background (Watkins, 1986).

In contrast, when students were predominantly identified using a range of procedures including teachers' opinions, the proportion of working class students fell and the proportion of middle class students increased, with representation of the former well below the levels indicated by tests of measured intelligence and representation of the latter well above.

More recent research by Watkins (ibid), in respect of students aged 15.5 to 16.5 years, provides evidence that on both widely employed tests of measured intelligence, and tests of reading comprehension and mathematical attainment, there were no significant differences between the sexes. Indeed, there is some slight evidence that female students out-performed their male counterparts on all aspects of the tests, including mathematical attainment. This situation prevailed following controls for social class and socio-economic status. In the same study Watkins found that on a subjective measure of assessment - impression marked English language essay scores - and following controls for measured intelligence and social class, female students out-performed their male counterparts by one full grade on a seven-point scale.

This might suggest that using similar subjective assessment measures, actually favours female students over the males, and to obtain a genuinely balanced pattern, largely objective measures provide an optimum solution.

Watkins (ibid) also found that in a sample exceeding 3,500 students, at the conclusion of their compulsory education in Britain, indices of attainment in all subjects at 16 plus (*ie* GCE 'O' Levels, CSE and others), female students out-performed their male counterparts, particularly in languages and social science disciplines, but also in the mathematics and science areas as well.

Other anecdotal evidence from universities and CAE'S in Australia and overseas strongly suggests that some ethnic minority groups are represented generally, and particularly in prestigious departments - medicine and law for example- disproportionately in far greater numbers than their representation in the wider community.

In the United States this situation has already given rise to concern with the apparently considerable over-representation (in terms of their population numbers within the wider American community as a whole) of Asian students within some areas and within the most prestigious institutions. While this information has not been substantiated by quantitative studies, it does suggest that minority groups are not necessarily disadvantaged by examinations and/or tests or other entry qualifications.

In view of the preceding, and an absence of definitive evidence which will sustain the notion that females, low socio-economic status and ethnic minority students are disadvantaged by examinations and/or tests, this generalised claim must be rejected as spurious. However, there is clearly cause for further investigation relating to the sub-sets within larger groups who appear to do particularly well or conversely particularly badly.

Thus, it is possible to advance a strong meritocratic argument in favour of objective system-wide tests or examinations on the grounds that such a system actually negates the potential prejudices of subjective assessment approaches, and facilitates the advancement of various minority groups (cf Douglas ibid).

Chapter 4 End Notes:

[1] Responses opposing external, system-wide examinations or tests were received from the Northern Territory Teachers Federation in respect of Towards the 90s Volumes 1 and 2, with remarks of a very similar vein appearing in the Territory Teacher Volume 14, No 8 published in December 1988

Chapter 5

Possible ways forward

Given arguments which continue to prevail about the use of all forms of assessment and reporting in respect of student performance and attainment, two questions arise:

'Is it possible to devise a system which successfully incorporates the best elements of system-wide, external examinations or tests and moderated school-based, continuous teacher assessment?'

Moreover:

'Is it possible to devise a system which successfully balances each of the components of assessment in such a way that all stakeholders are satisfied with a combination model, which does not require them to unduly compromise fundamental belief or value systems?'

If such an approach can be devised, not only should it incorporate the best features of external, system-wide testing or examining, the moderation of school assessment of students' work and teacher based continuous assessment, it should also preclude the worst features of each, thus avoiding over-subjectivity, narrowing the curriculum and de-contextualising student assessment.

Responses

It is in response to pressures of a very similar nature to those currently being experienced in Australia, that Britain has enacted its Education Reform Bill, which contains provision for a national curriculum and a new system of assessment for students.

It is also interesting to note the parallels between developments in Britain and those touched upon by the Minister for Employment, Education and Training in the recent paper Strengthening Australia's Schools (Canberra,1988).

Under the heading of A Common Curriculum Framework, Dawkins suggests:

"We need a curriculum that is relevant to our time and place in the world."

"Such a common curriculum framework could, for example, emphasise the need for higher general levels of literacy, numeracy and analytical skills across the nation. ..... Criteria for methods of assessing the achievement of curriculum objectives should be outlined. The framework should provide a guide to the best curriculum design and teaching practices.". (page 4)

This approach is extended in the discussion under the heading A Common Approach to Assessment, in which it states:

"There is a need for regular assessment of the effectiveness and standards of our schools. A common curriculum framework should be complemented by a common national approach to assessment. ..... A common approach to benchmarks for measuring student achievement, assessing school performance and public reporting on school-level education must be a further objective of a national effort.". (page 5)

Somewhat earlier, Beazley (ibid) made a number of recommendations concerning literacy and numeracy. These followed directly from the statement that:

"It is the perception of the Committee that the community expects schools to ensure high levels of literacy, and that the community would benefit if the Certificate of Secondary Education included a statement of each student's level of literacy.".

Accordingly Beazley (ibid) recommends that:

> "(i) The Board of Secondary Education define the literacy competencies students should have by the end of Years 10, 11 and 12.
>
> (ii) That schools decide whether students have adequate literacy skills, but the Board of Secondary Education, having responsibility for comparability of assessments, should provide assessment support materials and guide-lines for the assessments.
>
> (iii) That it should not be necessary for all students to undertake separate literacy courses in order to qualify as having reached an adequate standard in these skills.
>
> (iv) That the Board of Secondary Education undertake the development and implementation of these proposals."

In Britain, one of the Government's responses to concerns about national, system-wide student assessment was to establish a Task Group on Assessment and Testing. This group reported to the Secretary of State for Education and Science in mid-1988, and provided A digest for schools which summarised its recommendations.

Although the Task Group operated within the framework of Britain's proposed national curriculum (which incorporates 10 foundation subjects, including three core subjects in English, Mathematics and Science), it appears that much of its work can be applied to Australia, and to the Northern Territory.

It is important to note that the Task Group report A digest for schools recommends a number of strategies for overall student assessment. These include tests of various types; moderated school work; and individual teacher assessment.

These 'strands' will be aggregated to provide a much more complete profile of a student, than that likely to be available using any one method.

Critical sections of the Task Group report A digest for schools are cited below.

"The system of assessment proposed is in the main formative, in that its purpose is to provide information that will be of help when deciding how pupils' learning may be taken forward both individually and collectively. **For 16 year olds who are at the end of the statutory period of education, the system is also summative, providing a comprehensive picture of the overall achievements of a pupil.**" (page 6-7)

"If external tests are to provide the touchstone, they have to be carefully designed so that their results can be trusted. ..... The art of constructing good assessment tasks is to exploit all (*ie a wide range of*) possibilities. ..... Paper and pencil tests have an important part to play .....".(page 11)

"**The national system should employ tests for which a wide range of modes of presentation, operation and response should be used so that each may be valid in relation to the attainment targets assessed.** These particular tests should be called 'standard assessment tasks' and they should be so designed that flexibility of form and use is allowed wherever this can be consistent with national comparability of results." (page 21)

Before considering further, general recommendations made by the Task Group, it may be appropriate to recall that they were addressing assessment strategies for students at the ages of 7, 11, 13 and 16.

"Any system of assessment should satisfy certain general criteria. For the purpose of national assessment we give priority to the following four criteria:

- the assessment results should give direct information about pupils' achievement in relation to objectives: they should be criterion referenced;

- the results should provide a basis for decisions about pupils' further learning needs: they should be formative;

- the scales or grades should be capable of comparison across classes and schools, if teachers, pupils and parents are to share a common language and common standards: so the assessments should be calibrated or moderated; and

- the ways in which criteria and scales are set up and used should relate to expected routes of educational development, giving some continuity to a pupil's assessment at different ages: the assessments should relate to progression."

Thus the Task Group report A digest for schools suggests the establishment of an important range of principles and procedures for student assessment and reporting (this latter item is not discussed here).

In adopting the widest possible combination of approaches to the assessment of student performance and attainment, the Task Group have been clearly conscious of the needs of stakeholders in the education system - students, parents, teachers, employers and others - and have accordingly suggested ways of meeting each group's needs while providing students with an optimum system of assessment.

It is possible that approaches of this type may also find favour in Australia and in the Northern Territory, and that the sometimes acrimonious debate about the assessment of student performance and attainment, may cease to be an issue constantly about to erupt from any one quarter.

Bibliography and References

Australia Council for Educational Research (1981) Newsletter, Vol 43

Adelman Clifford. (1984) The Standardised Test Scores of College Graduates, 1964 - 1982. Study Group on Conditions of Excellence in American Higher Education.

Beazley K E. (1984) Education in Western Australia - A Report of the Committee of Inquiry, Education Department of Western Australia, Perth.

Bresser J. (1987) National Testing. Australian Primary Principals' Association Newsletter, Volume 15, No 1 Sept/Oct 1987

Bourke S F. (1979) Community Expectations of Numeracy in Schools, Australian Journal of Education, 23, 2: pp 142-152

Bourke S F. (1982) A Case for Testing. Curriculum Perspectives, Vol 2, No 3, October 1982.

Bourke S F and Keeves J P. (eds) (1977) The Mastery of Literacy and Numeracy, Australian Government Publishing Service, Canberra.

Bourke S F, Mills J M, Stanyon J and Holzer F. (1981) Performance in Literacy and Numeracy, Australian Government Publishing Service, Canberra

Boyer E L. (1983) Carnegie Foundation High School: A Report on Secondary Education in America. Harper and Row, New York.

Cavanagh D. (1982) To Test Or Not To Test: Can There Be Any Question?, Curriculum Perspectives, Vol 2, No 3, October 1982.

Chambers 20th Century Dictionary. (1983) W R Chambers, Edinburgh.

Commission on Pre-college Education in Mathematics, Science and Technology. (1983) Educating Americans for the 21st Century

Concise Oxford Dictionary. (1976) Oxford University Press, Oxon,

Davidson, G S (Chairman). (1981) Preparation for the Workforce. Senate Standing Committee on Education and the Arts, AGPS, Canberra

Dawkins J. (1988) Strengthening Australia's Schools. Ministerial Statement. Department of Employment, Education and Training, Canberra, ACT.

Dockerell W. (1984) Assessing What Teachers Teach. International Association for Educational Assessment, 10th International Conference, Perth, June 1984

Douglas J W B. (1964) The Home and the School; Panther Books, London.

Dreher M J and Singer H. (1984) Making Standardised Tests Work for You. Principal March 1984.

Education Department of Tasmania. (1988) 1987 Survey of Basic Reading Skills of 14-Year-Old Tasmanian Students, Hobart, November, 1988

Education Department of Western Australia. (1977) Policy from the Director-General's Office. No 6, August.

Fogelman Ken. (1984) Problems in Comparing Examination Attainment in Selective and Comprehensive Secondary Schools. Oxford Review of Education, Volume 10, No 1.

Goldstein Harvey. (1984) The Methodology of School Comparisons, Oxford Review of Education, Volume 10, No 1.

Gray J, Jesson D and Jones J. (1984) Predicting Differences in Examination Results Between Local Education Authorities: does school organisation matter. Oxford Review of Education, Volume 10, No 1.

Grosvenor Rod. (1987) The Tasmanian Certificate of Education. Prepared for the National Seminar - Reports and Records of Achievement for School Leavers, Sydney, NSW.

Harlen, W (1979) Accountability that is of benefit to Schools. Journal of Curriculum Studies, 11, 4, pp 127-140

Heim Alice. (1970) the Appraisal of Intelligence. NFER, London.

Heim Alice. (1970) Intelligence and Personality. Penguin, London.

Husen T. (1972) Social Background and Educational Career, OECD, Paris.

Husen T. (1975) Social Influence and Educational Attainment, OECD, Paris.

Irby Alice. (1985) Uses of the Outcomes of assessment - the National Context. International Association for Educational Assessment, 11th Annual Conference, Oxford.

Lipset S M. (1972) Social Mobility and Educational Opportunity. Public Interest Journal, Fall, pages 90-108.

Macintosh H G. (1983) Public Examinations and the Development of Reasoning Skills. International Association for Educational Assessment, 11th Annual Conference, Oxford

Marks J, Cox C, and Pomian-Srzednicki. (1983) Standards in English Schools, National Council for Educational Standards, London.

Mathematics Counts. (1982) Report of the Committee of Inquiry Into the Teaching of Mathematics in Schools under the Chairmanship of Dr W H Cockcroft, London, HMSO, January.

Mortimore J and Blackstone T.(1982) Disadvantage and Education: Studies in Deprivation and Disadvantage. Heinemann Educational Books, London.

National Commission of Excellence in Education. (1983) A Nation At Risk.

National Curriculum. (1988) Task Group on Assessment and Testing Report: A digest for schools. Department of Education and Science and the Welsh Office, London.

Northern Territory Teachers Federation. (1987) Submission to the Education Advisory Council (EAC) in response to Towards the 90s: Volume 1, Darwin.

Northern Territory Teachers Federation. (1988) The Territory Teacher, Volume 14, No 8, page 5, December, Darwin.

O'Loughlin M, J, Bardsley and J Bird. (1984) Evaluating a School's Performance in External Examinations. 10 Annual Conference of the International Association for Educational Assessment, Perth, June 1984.

Popham W J and Husek T R. (1969) Implications for Criterion Referenced Measurement. Journal of Educational Measurement, 6, 1: pp 1-9

Robbins (1973) A Report into the Composition of Higher Education in Britain, HMSO, London.

Roget's Thesaurus. (1980) 16th Impression, Longman Publishers, London.

Secondary Education Authority of Western Australia. (1987) Third Annual Report 1986-1987, Perth

Skills for Australia.(1987) Ministerial Statement by J S Dawkins and A C Holding, Department of Employment, Education and Training, Canberra.

Stetz F P and Beck M D. (1980) Teachers Oppose Halting All Standardised Testing. Phi Delta Kappan, #61, p648.

Task Force on Economic Growth. (1983) Action for Excellence. Washington DC

Taylor Gordon. (1988) Is the ability of our students to handle the English language declining. Education Victoria, Volume 3, Number 4.

The Australian: 6 January 1988

The Australian: 14 January 1988

The Australian: 2 December 1988

The Northern Territory News: 28 January 1988

The College Board. (1983) Academic Preparation for College: What Students Need to Know and Be Able to Do, New York.

The Times Educational Supplement.(1987) The Quest for Standards; G Howson and R Murphy. London, 16 October 1987

The Times Educational Supplement. (1987) Defence not the best attack. A Syred. London, 13 November 1987

The National Assessment of Educational Progress. (1975) U S Government Printing Office, Washington DC

Towards the 90s: Volume 1: Excellence, Accountability and Devolution.(1987) Education Department of the Northern Territory Discussion Paper, Darwin.

Towards the 90s: Volume 2. (1988) Education Department of the Northern Territory Discussion Paper, Darwin.

Twentieth Century Fund. (1983) Making the Grade: Task Force on Federal Elementary and Secondary Education Policy, The Fund, New York.

Vernon P E. (1960) Intelligence and Attainment Tests. University of London Press, London.

Vernon P E. (1965) The Measurement of Abilities. University of London Press, London.

Vernon P E. (1969) Intelligence and Cultural Environment. Methuen and Co, London.

Victorian Curriculum and Assessment Board. (1987) Newsletter - Special Edition. VCAB, Melbourne, October 1987.

Victorian Curriculum and Assessment Board. (1988) Assessment for the Victorian Certificate of Education. VCAB, Melbourne,March, 1988.

Victorian Curriculum and Assessment Board. (1988) Common Assessment Tasks (CAT's); VCAB, Melbourne, June, 1988.

Victorian Curriculum and Assessment Board. (1988) Newsletter VCAB, Melbourne, October 1987.

Watkins R. (1981) Language, Intelligence and Expectation in the Education of Ethnic Minority and Socially Disadvantaged Students. Unpublished PDESL Dissertation, University of Leeds.

Watkins R. (1986) Factors involved in language performance and their relationships with attainment. Unpublished PhD Thesis, University of Leeds.